

DECLARATION

I, Daneen Milam, do hereby declare the following to be true and correct, under penalty of perjury:

CREDENTIALS

1. I received my Ph.D. in Educational Psychology from Texas A & M University in 1982. I am a licensed psychologist and certified as a Health Service Provider in the State of Texas. My license number is 2-2661.
2. I am board certified in clinical neuropsychology by the American Board of Professional Neuropsychology. I practice clinical neuropsychology in San Antonio, Texas, where I was the director of an assessment center for more than eighteen years. I have served for five years on the editorial board of Archives of Clinical Neuropsychology, the official scholarly journal of the National Academy of Neuropsychology.
3. Much of my current clinical practice is in the area of forensic neuropsychology. I have conducted more than fifty (50) examinations of criminal defendants. I have testified in excess of seventy-five (75) times as an expert witness in the state courts of Texas for the Texas Department of Family and Protective Services, Child Protective Services, and more than twenty (20) times as an expert witness in state murder trials. My curriculum vitae is attached to this declaration.
4. In January 2010, I testified at Mr. Long's state habeas hearing.

PURPOSE OF THIS DECLARATION

5. At the request of current federal habeas counsel for Mr. Long, I was asked to address two specific questions, which were not completely addressed during Mr. Long's state habeas hearing. The first question was whether Mr. Long could be diagnosed as mentally retarded with a high level of certainty. Secondly, was it possible to rule out malingering with a high level of certainty? To answer both of these questions a comparison was undertaken by

evaluating Mr. Long's current level of mental retardation, also called intellectual disability, and a current measure of malingering and then comparing these scores with obtained scores from prior evaluations. Both the current measures of intelligence and the 2009 scores are based upon normative samples developed by the Wechsler Adult Intelligence Scales-Fourth Edition. The malingering indices are both based upon the Advanced Clinical Solutions, using the normative data of mental retarded individuals along with an item by item comparison of past performances and current performance. This ipsative comparison, comparing Mr. Long's present performance and Mr. Long's prior performance would be difficult if not impossible to malingering and is a strong indicator of his honesty and effort in his evaluation of mental retardation. Secondly, over the last four years research in the area of malingering has become quite pronounced as to the inadequacy of comparing the mentally retarded to base rates achieved by other populations such as traumatic brain injury and the emotionally impaired. Since the Advanced Clinical Solutions, (ACS) which had been designed to address the inadequacies of former measures of malingering, had only been on the market for a matter of weeks at the time of his state habeas hearing, a request for a measure by measure examination of his performance on these scales in 2009 to be compared to his performance in 2013 was critical to ensure a fair and honest appraisal of Mr. Long's true level of functioning as well as to address the issue of whether he was or was not malingering to avoid execution. A comparison of the 2009 and 2013 scores are provided in Addendum A.

RESOURCES

6. My analysis for these questions involved the review of multiple sets of records related to Mr. Long's social history and his trial, including trial testimony, my own notes as an expert witness, and my notes from a 2009 interview of Mr. Long. I also reviewed testimony given by Dr. Randall Price and articles he has written in the area of forensic evaluations of malingering. I reviewed the most current forensic data on malingering, as well as recent studies on the base rates of scores of the mentally retarded as well studies of sensitivity and specificity among the mentally retarded. I drove to the Polansky Unit in Livingston and evaluated Mr. Long over a two day period to obtain current scores to compare to prior evaluations and to obtain a current measure of

malinger. These scores were used to complete an ipsative comparison (comparing Mr. Long to himself over time.)

HISTORY OF MALINGERING INSTRUMENTS

7. The assessment of intellectual potential began during World War II and the Wechsler series are the primary instrument used to measure adult and adolescent intelligence. The initial version of the Wechsler Adult Intelligence Test was published in 1955 and the definition of intelligence was defined as "The global capacity of a person to act purposefully, to think rationally, and to deal effectively with his or her environment." (WAIS Manual). The latest version of the instrument was published in 2008 and dramatically departed from earlier versions that were reported as Verbal IQ, Performance IQ and Full Scale IQ. There are ten core subtests and five supplemental subtests on the WAIS-IV version and measures of intellectual potential are reported as indices. The structure of the Wechsler Series has been the subject of numerous studies. The new WAIS-IV was standardized on 2200 people and the newest version of the Wechsler Memory Scale was revised and updated at the same time on the same sample of individuals who comprise the standard sample. Research indicates that the WAIS-IV with its new configuration of four index scores, along with special scores to identify malingering, often called suboptimal effort, is a more appropriate and better test than the previous editions and a much more reliable measure of validity.
8. At the same time measures of brain integrity were being developed to understand brain-behavior relationships. The goal of these instruments was to ascertain how lesions in specific areas of the brain could change behavior in specific ways. The primary goal was to predict from a given behavior the site of a lesion and predict from a lesion the probable behavioral abnormality. The study of brain lesions in humans and animals has yielded more information about brain-behavior relationships than any other approach and was the forerunner of modern computerized tomography (CT Scans), and Positron emission tomography (PET Scans.) Malingering scales were developed from embedded items across the instruments used in a Neuropsychological evaluation.

9. Initially, there was a fundamental assumption that the individuals undergoing a Neuropsychological Evaluation and the practitioners shared a common goal in the assessment process. However, as the shift from research of brain behavior relationships to forensic evaluations occurred and the potential for financial remuneration following demonstration of cognitive impairment increased, the temptation to feign dysfunction became more and more prevalent. By the 1970's and 1980's it was estimated that 40% of all individuals being evaluated for workers compensation, or personal injury litigants were either exaggerating their symptoms, making insufficient effort, or feigning symptoms (Slick, 1999). There was a significant need for instruments to identify true brain damage from malingering. The common definition of malingering was generally held to be "the intentional production of false or grossly exaggerated symptoms that are motivated by external incentives." (Marshal and Happe, 2007). There was a growing consensus in the field of Neuropsychology that the assessment of malingering, often called suboptimal effort or feigned cognitive impairment was a fundamental component of forensic evaluations which should be assessed with each Neuropsychological Evaluation. These tests, collectively, are called effort tests.

OVERVIEW of MALINGERING INSTRUMENTS

10. By 2000 several effort tests were available to detect potential malingers. For the most part, the efforts tests would report sensitivity rates (it measures what it proposes to measure and identifies individual with that trait). The most commonly administered effort tests, which were all test of memory, were:

TOMM:	Test of Memory Malingering
VIP:	The Validity Indicator Profile
FIT:	The Rey 15 Item Test
DCT:	The Dot Counting Test
LMT:	The Letter Memory Test
WMT:	The Word Memory Test
MSVT:	The Medical Symptom Validity Test

These instruments were clearly more effective than a standalone clinical

Interview or an opinion based upon clinical experience. Base rates (how often something is seen in a specific population) was developed for a wide variety of subgroups. Cut off scores were developed by disorder, by age, by education, and most often brain injury. Many studies were undertaken but people with intellectual disability were believed to perform like the brain damaged and when the individuals with intellectual disability were evaluated the cut off scores for the brain damaged samples were used. Very few studies of individuals with low scores on measures of potential were undertaken at all. Franzer and Martin (1996) took the position that intelligence played little or no role in the ability to malingering, and this was an opinion held by many. This subset of the population was rarely tested while large groups of people with traumatic brain injury filled the literature. It was not until 2002, with the Atkins vs. Virginia Supreme Court decision that mental retardation and the malingering of mental retardation became a major issue and unfortunately the instruments used to document suboptimal effort had never been evaluated to ensure the psychometric properties could be extended to a population of mild to moderate mentally retarded.

11. From 2003 to 2006, as more and more inmates on death rows across the country were evaluated for intellectual disabilities, a small but vocal group began to report that the intellectually challenged tended to score like the malingering with approximately the same cut off scores, thus producing an unacceptable rate of false positives (Victor and Boone (2007)). One of these reporters was Dr. Price who coauthored an article in The Journal of Forensic Neuropsychology. He wrote in 2003 "In spite of the psychometric sophistication of the malingering measures presented, their use is to detect malingering brain injury or even a specific aspect of neuropsychological functioning such as memory. None were designed to detect malingering mental retardation." Yet the instruments he profiled in his article were measures he used to determine Mr. Long was not mentally retarded when his own testing indicated an overall intellectual ability in the mild mentally retarded range. Due to the scores on short paper and pencil tasks and Mr. Long's verbal statements that he had read many books, Dr. Price came to believe the defendant was malingering his intellectual disability and never addressed the flaws in his own instruments while minimizing the high standards of validity of the ACS. This rush to judgment allowed him to accept Mr. Long's statement as true that he read many books in the Grisham series

without giving Mr. Long a measure of reading comprehension. Had he given such an instrument, he would have found that Mr. Long could phonically sound out words but cannot read for meaning. On a formal reading comprehension task, Mr. Long was found to be reading at the beginning third grade level which was commensurate with his measures of intellectual potential.

12. By 2008, the standard of care had evolved to a commonly held belief that many instruments used to measure effort were at risk of false identification of intellectually disabled individuals. It was commonly held to be true that many intellectually disabled individuals were being falsely accused of not putting forth maximum effort when they had honestly tried their best. It was not that the intellectually disabled could not be evaluated for suboptimal effort but they needed (1) adequate norming samples of mentally challenged people to be compared against (2) cut off scores that reflected the performance of peers of same level potential, (3) at least three different types of instruments, such as forced choice and embedded designs. All the while inmates on death rows across the country were being evaluated for mental retardation using antiquated measures of malingering while life and death decisions were being made. For the courts to allow practitioners to testify that an individual is not mentally retarded when they score in the mentally retarded range on measures of intelligence but score in the malingering range on forced choice and memory tests is not close to the accepted standard of care in the assessment community and is an egregious error. It is neither fair nor does it serve the cause of justice. At this time, due to the high number of false positives on standard malingering instruments, the Social Security Service DDS has formally discouraged the use of formal measures of effort to determine eligibility.

THE ADVANCED CLINICAL SOLUTIONS

13. The Wechsler Adult Intelligence Scales -IV was introduced in 2008 and the Wechsler Memory Scales- IV was introduced in 2009 along with the Advanced Clinical Solutions (ACS). The WMS-IV is a battery of subtests that are frequently administered in a comprehensive measure of memory. The

WMS-IV contributes three scores to the ACS while the WAIS-IV contributes one test and a new forced choice test completes the five subtest battery. Since the ACS is a subtest of the WAIS-4 and WMS-4, it is normed on 2,200 individuals. These are well developed standardized instruments. After reviewing the literature it was clear to the test design team of the WAIS-4 that a valid measure of malingering with strong test integrity and sound validity and reliability was critical. While it was true that there were significant problems with the instruments on the market used to decide life and death decisions, it was also true that without a sound measure of effort put forth by prisoners the results obtained would be of questionable validity and accuracy. There was an urgent need for an empirically derived prediction model to detect suboptimal effort by individuals incarcerated that had a vested interest to appear incapable of making rational choices. Several of these problems were addressed in the ACS.

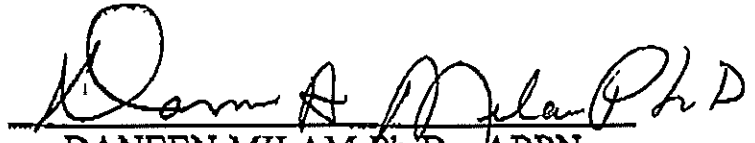
14. The expected profile for the individual with intellectual deficits is to have difficulties with slowed thinking, reduced processing speed, memory deficits and impaired attention. Memory can be divided into higher level cognitive functioning while lower order cognitive functions are exhibited as attention and processing speed deficits and all must be measured to differentiate between mental retardation and suboptimal effort. None of the measures of malingering that were commonly used are adequate for this task and the authors of these instruments have uniformly acknowledged this in their operating manuals.
15. To be mentally retarded, an individual must meet three prongs:
 - a. Significantly sub average general intellectual functioning
 - b. Significant limitations in adaptive functioning in at least two skill areas; communication, self-care, home living social/interpersonal skills, use of community resources, self-direction, functional academic skills, work, leisure, health and safety.
 - c. The onset must occur before the age of 18
16. Malingering may be defined as the presence of clear and compelling data indicating volitional exaggeration or fabrication of cognitive dysfunction

without plausible evidence of alternative explanations for such a cognitive profile (Tombaugh, 1996). To diagnose malingering there must be:

- a. Element of intentional production of false or grossly exaggerated symptoms that are motivated by external incentives.
- b. Intentionally put forth suboptimal effort during an evaluation.

17. Using scales developed to update the newest version of the WAIS-IV and WMS-IV, along with norms developed from known samples of individuals documented to be intellectually disabled, Mr. Long was found to be functioning in the mentally retarded range. Mr. Long's scores are replicated across four assessments. He also exhibited a severely limited ability to care for himself and there is significant antidotal evidence of this fact available which meets the criteria of poor adaptive functioning.
18. It is my opinion that the foregoing description of the malingering instruments given, taken with Dr. Price's description of the ACS as having no documented reliability and reliability and ignoring the widespread adoption of a specialized perspective in which standardized testing is paired with measures specifically designed to detect feigning using base rates and cut off scores, a balanced assessment of Mr. Long's level of personal moral culpability has not been addressed.
19. In 2010 an effort scale could not be computed from Dr. Price's work because Dr. Price did not give enough of the subtests needed to develop a malingering scale. In 2013 Mr. Long produced results that were compared with individuals who were well documented as mentally retarded. His scores on his five subtests, scored at the 5% base rate, produced three scores that were at or above the 5% cut off score. When compared with the norming sample, 38% of the documented mentally retarded produced scores commensurate with Mr. Long's performance. Therefore, a rate of more than one in three would have been a false positive. By looking at all five subtests, the error rate would have been zero.
20. Due to Mr. Long's consistency across all evaluations, his pattern of responses on measures of malingering, and his lifelong functional limitations, it is my opinion that Mr. Long is mentally retarded.

I, Daneen A. Milam, declare under penalty of perjury that the foregoing is true and correct. Executed on this 26 day of February, 2013 in San Antonio, Texas.


DANEEN MILAM Ph.D., ABPN

ADDENDUM A

A raw score of each subtest from three administrations of the Wechsler series was undertaken to address the consistency of effort to partially answer the question of malingering. The 2006 testing was undertaken using the Wechsler Adult Intelligence Scales-III and the 2009 and 2013 testing was undertaken using the Wechsler Adult Intelligence Scale -IV.

	2006	2009 Dr. Price	2013 Dr. Milam
Full Scale IQ	62	64	55
Verbal Composite	66	70	56
Perceptual Composite	64	73	63
Working Memory	--	71	66
Processing Speed	--	62	62
General Ability	--	69	69

SUBTESTS: Reported in Raw Scores

Block Design	20	24	16* Significantly Different
Similarities	4	13	4* Significantly Different
Digit Span	11	17	16
Matrix	---	6	7
Vocabulary	5	22	10* Significantly Different
Arithmetic	8	8	7
Symbol Search	---	12	11
Visual Puzzles	---	8	6
Information	5	4	3
Coding	40	28	26
Comprehension	---	---	5
Figure Weights	---	---	4
Picture Completion	---	---	2

A second measure of intelligence was given by Dr. Price and Dr. Milam

Reynolds Intellectual Assessment Scales

CIX is a Full Scale IQ	63	63
VIX is a measure of Verbal Skills	71	71
NIX is a measure of visual perceptual skills	65	64
CMX is a measure of memory skills	---	65

DISCUSSION OF SCORES:

The three evaluations of Mr. Long's intellectual potential took place over a six and a half year period and Mr. Long was evaluated by three separate examiners. For the most part, his answers, scores and indices remained the same with three subtests that warrant discussion.

- (1) Using an item by item comparison, Mr. Long got the same number of items right on Dr. Price's evaluation of the subtest of Block Design and his most current one. He did not get credit for those two items due to failure to meet the time cut off. This score represents a minor fluctuation of performance that is common in a measure of potential and/or a failure to adhere strictly to time limitations on individual items.
- (2) His performance on a subtest that requires abstract thinking was somewhat more troubling. The question(s) on this subtest requires the subject to figure out how things are related to one another. An answer can be scored zero, one or two.

When evaluated in 2006 the question was "In what way is a dog and a lion the same?" and he responded "They are not the same, a lion is dangerous and a dog is friendly." This answer produced a score of zero.

When he was evaluated by this examiner in January of 2013, the question was "In what way is a horse and tiger the same?" He responded "They are not the same. This produced a score of zero.

When Mr. Long was evaluated by Dr. Price in 2009, the question was "In what way is a horse and tiger the same? Dr. Price stated he said "animal" and gave him credit of 2 points which allowed him to continue and thus produced a significantly higher score on this measure. Therefore, in 2006 Mr. Long received a score of 4 on this subtest and in 2013 he received a score of 4 on this subtest but in 2009 Mr. Long received a score of 13 on this subtest, along with an exceptionally high score on a measure of vocabulary. In 2006 and 2013 Mr. Long had an exceptionally limited vocabulary but when evaluated by Dr. Price in 2009 he could define words such as assemble, consume, and ponder. Taken together, these anomalies increased his indices from (56-63-66) to (70-73-71). Even with these differences, Mr. Long's Full Scale IQ score across the years and across examiners remained in the mentally retarded range (2006= 62, 2009=64, 2013=55). The Full Scale IQ is considered to be the most accurate measure of intellectual integrity.

Therefore, using these three evaluations of consistency, Mr. Long's scores remained surprisingly the same. Even an individual with a superior level of intellectual potential could not retain the same response pattern, and basically the same answers across six and a half years, over two forms of the Wechsler series, and three different examiners. The only way an individual could keep this level of consistency would be answering honestly and making an effort.

Addendum B: Measures of Malingering

The Test of Memory and Malingering (TOMM) was given in both the 2006 and the 2009 evaluation.

	2006	2009
Trial One	32	34
Trial Two	34	38
Retention Trial	33	33

In both instances, Mr. Long's scores did not meet the number of correctly answered responses to pass the cut off scores for malingering. However, his scores for retention of information were remarkably consistent.

Advanced Clinical Solutions: Base Rates at the 10% Cut off Scores

	One Tests	Two Tests	Three Tests	Four Tests	Five Tests
Simulators	40%	24%	14%	10%	
Intellectually Disabled	38%	13%	3%	3%	0%

Simulators were individuals with average to above average intelligence who were asked to simulate mental retardation and then were told that if they succeeded in deceiving the examiner they would get a cash reward.

Intellectually Disabled base rates were developed from individuals who have been diagnosed as mentally retarded for several years and were receiving state help for their handicap.

Looking at percentages of cases in these two groups, using a 10% cut off score, with a score below the cut off score to be labeled as malingering:

Simulators scored 64% of the time below the cut off score of one test.
 Simulators scored 36% of the time below the cut off score on two tests.
 Simulators scored 20% of the time below cut off score on three tests.

Simulators scored 10% of the time below cut off score on four tests.
Simulators scored 5% of the time below cut off score on five tests.

The intellectually disabled scored 63% of the time below cut off score on one test.

The intellectually disabled scored 31% of the time below the cut off score on two tests.

The intellectually disabled scored 16% of the time below the cut off score on three tests.

The intellectually disabled scored 6% of the time below the cut off score on four tests.

The intellectually disabled scored 0% of the time below the cut off score on five tests.

What can be understood from this data, 63% of well documented intellectually disabled individuals who took malingering tests failed at least one measure of malingering. Thirty one percent of these individuals failed two tests and sixteen percent of these individuals failed three tests. These individuals would be labeled as malingering and they would have been answering honestly with good effort. Sixteen percent is too high of a number of false positives when life and death decisions are being made. However, using five well normed tests and compared with other ability equal individuals, the false positives fall to zero.

The Social Security Administration, based upon the emerging statistics in the field have "strongly discouraged" the use of effort tests in the evaluation of the disabled.

Addendum C: Bibliography

- Arvey, R.D. et. Al (1994, December 13th). Mainstream science on intelligence. The Wall Street Journal. pA18.
- Atkins V. Virginia, 536 U.S. 304 (2002).
- Fredrick, R.I. (1997) Validity Indicator Profile Manual, NCS Assessments, Minnetonka, MN.
- Green P. (2004). Test Manual for The Medical Symptom Validity Test. Edmonton: Green Publishing.
- Hurley, Kollen E. and Durn, William P.; Assessment instruments measuring malingering used with individual who have mental retardation; Mental Retardation, Vol 44, Number 2; pp 112-119. April (2006).
- Marshall, P and Happe, M. (2007) The performance of individuals with mental retardation on cognitive tests assessing effort and motivation. Clinical Neuropsychologist, 21, pp. 826-840.
- Reynolds, C.R., Price, J.R., and Niland, J. Journal of Forensic Neuropsychology.
- Applications of Neuropsychology in capital felony (Death Penalty) defense, pp. 89-122.
- Report on the Task Force on Mental Disability and the Death Penalty (2005).
- Slick, D.J. Sherman, E.M.S. and Iverson, G.L., (1999) Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. Clinical Neuropsychologist, 13 (4), 545-561.
- Tombaugh, T. N., (1996). The test of memory malingering. Toronto Canada. Multi-Health Systems.
- Victor, T. L. and Boone, R.D. (2007). Identification of feigned mental retardation.
- In K.B. Boone (Ed.) Assessment of feigned cognitive impairment: A neuropsychological perspective. (pp. 310-345). New York: Guilford.
- Wechsler, D. (2008). Wechsler Adult Intelligence Test, 4th edition. (WAIS-4). San Antonio, Texas: Psychological Corporation.